PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/62, 15/85, 5/10, G01N 33/68

A1

(11) International Publication Number: **WO 98/49325**

(43) International Publication Date: 5 November 1998 (05.11.98)

(21) International Application Number: PCT/CA98/00385

(22) International Filing Date: 24 April 1998 (24.04.98)

(30) Priority Data:
60/044,911 28 April 1997 (28.04.97) US

(71) Applicant *(for all designated States except US)*: McGILL UNIVERSITY [CA/CA]; 845 Sherbrooke Street West, Montréal, Québec H3A 1B1 (CA).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: STOCHAJ, Ursula [DE/CA]; 1100 Dr. Penfield Street #601, Montréal, Québec H3A 1A8 (CA).

(74) Agent: COTÉ, France; Swabey Ogilvy Renault, Suite 1600, 1981 McGill College Avenue, Montréal, Québec H3A 2Y3 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: REPORTER FUSION PROTEINS, EXPRESSION VECTORS AND TRANSFECTED CELL LINES THEREOF FOR THE ANALYSIS OF NUCLEAR TRANSPORT

(57) Abstract

The present invention relates to reporter fusion proteins which allow the monitoring of facilitated transport and passive diffusion of proteins across the nuclear envelope in mammalian cells, cloning and expression vectors encoding these reporter proteins and transfected cell lines thereof. More particularly, the present invention relates to genes encoding reporter proteins to monitor transport and passive diffusion of proteins across the nuclear envelope in mammalian cells, which comprises a protein carrying a nuclear localization sequence (NLS) fused to at least one copy of a reporter protein, wherein the reporter protein allows for the visualization of the reporter fusion protein.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

REPORTER FUSION PROTEINS, EXPRESSION VECTORS AND TRANSFECTED CELL LINES THEREOF FOR THE ANALYSIS OF NUCLEAR TRANSPORT

BACKGROUND OF THE INVENTION

(a) Field of the Invention

The invention relates to reporter fusion proteins which allow the monitoring of facilitated transport and passive diffusion of proteins across the nuclear envelope in mammalian cells, cloning and expression vectors encoding these reporter proteins and transfected cell lines thereof.

(b) Description of Prior Art

The correct localization of proteins in eukaryotes involves complex targeting reactions. These processes ensure that each cellular compartment contains an unique set of macromolecules required for the specialized function of the organelle. Proteins that are destined for the nucleus carry specific targeting sequences, termed nuclear localization sequences (NLSs). The NLS is recognized by cytoplasmic NLS-receptors that dock at the nuclear pore complex which subsequently catalyzes the translocation across the nuclear envelope. Additional soluble factors are required for these transport reactions.

The nuclear pore complex mediates also diffusion, and proteins of molecular masses of 40kD or less can enter the nucleus via passive diffusion. By contrast, proteins that exceed the size of the diffusion channel enter the nucleus by a facilitated process that requires energy and a NLS. Polypeptides larger than 70kD are excluded from nuclei if they do not carry a functional NLS. NLSs can be divided into several classes. The NLS present in SV40 T-antigen is an example for a "simple" NLS containing a stretch of basic amino acid residues. A second class of NLSs contains "bipartite" sequences that carry two clusters of posi-

tively charged amino acid residues that are separated by a spacer region, such as the NLS of *Xenopus laevis* nucleoplasmin. Another type of targeting sequences includes complex NLSs which do not fit into the classes of simple or bipartite signals.

So far, most experimental systems for the analysis of protein import in higher eukaryotes were based on *in vitro* reconstitution of transport or the injection of nuclear substrates into single cells. These experiments are tedious and time-consuming. A caveat of some of the reporter proteins used in these studies is their post-translational modification such as phosphorylation which can modulate nuclear traffic.

It would be highly desirable to be provided with improved tools to monitor nuclear transport of proteins *in vivo* and *in vitro*.

SUMMARY OF THE INVENTION

For the purpose of the present invention the following abbreviations are defined below.

ADPP	ala-asp-pro-pro
APGP	ala-pro-gly-pro
BSA	bovine serum albumin
DAPI	4',6'-diamidino-2-phenylindole
FACS	Fluorescence activated cell sorting
GFP	*Aequorea victoria* green fluorescent protein
GNSP	gly-asn-ser-pro
NLS	nuclear localization sequence
NP-NLS	NLS of *Xenopus* laevis nucleoplasmin
PBS	phosphate buffered saline
pKS^+	plasmid pBluescript KS^+
SV40-NLS	NLS of SV40 T-antigen
SV40*-NLS	mutant form of SV40-NLS
SV40inv-NLS	inverse sequence of SV40-NLS.

To analyze nucleocytoplasmic traffic in growing cells, an experimental system which is based on the inducible synthesis of fluorescent reporter proteins was developed. Reporter proteins of various sizes described here carry different NLSs which are fused to the *Aequorea victoria* green fluorescent protein (GFP). These substrates were expressed and localized in mammalian culture cells. In addition to the analysis of facilitated protein transport into nuclei, assays to study passive diffusion from the nucleus into the cytoplasm were described. To facilitate the generation of gene fusions that encode different mutants of GFP with or without a nuclear localization sequence several new cloning vectors were generated.

Moreover, in accordance with the present invention, a stable cell line that produces a nuclear reporter protein under control of an inducible promoter was analyzed. Upon induction, single cells of this clone synthesize similar amounts of the reporter protein. In accordance with the present invention, such cell lines offer the advantage that large numbers of cells can be tested simultaneously under identical conditions. Additional stable cell lines are currently generated for more detailed analyses of transport and diffusion across the nuclear envelope. Taken together, novel experimental tools to investigate nucleocytoplasmic transport and diffusion in intact mammalian cells were described.

In addition to the analysis of nuclear traffic in growing cells, tagged nuclear transport substrates carrying various numbers of GFP to study nucleocytoplasmic traffic *in vitro* were also generated. These substrates can be expressed in *Escherichia coli* and purified by metal affinity-chromatography. The purified substrates are currently tested with semi-permeabilized

HeLa cells and with yeast spheroplasts to demonstrate that these proteins are bona fide substrates for nuclear transport.

In accordance with the present invention there is provided genes encoding reporter proteins to monitor transport and passive diffusion of proteins across the nuclear envelope in mammalian cells, which comprises a protein carrying a nuclear localization sequence (NLS) fused to at least one copy of a reporter protein, wherein the reporter protein allows for the visualization of the reporter fusion protein.

In accordance with one embodiment of the present invention, the NLS derived from SV40 T-antigen is connected to the reporter protein via a linker sequence encoding small amino acid residues. The NLS may also be derived from *Xenopus laevis* nucleoplasmin or present in the C-terminal portion of nucleoplasmin.

In accordance with one embodiment of the present invention, there is provided an expression vector to transiently and stably express reporter genes of the present invention, which comprises a gene encoding a protein carrying a nuclear localization sequence (NLS) fused to at least one copy of a reporter protein, wherein expression of the gene is under the control of an inducible promoter.

The NLS may be connected to the reporter protein via a linker sequence encoding small amino acid residues.

In accordance with another embodiment of the present invention, the reporter protein may be any fluorescent protein such as *Aequorea victoria* green fluorescent protein (GFP).

In accordance with one embodiment of the present invention, there is provided, a cell line stably expressing a nuclear reporter protein to monitor trans-

port and passive diffusion of proteins across the nuclear envelope, which comprises a cell line transfected by the expression vector of the present invention.

The NLS may be fused to the reporter protein by a linker encoding small amino acid residue.

The cell lines may be selected from the group consisting of HeLa, COS and CHO.

In accordance with another embodiment of the present invention, there is provided an *in vivo* assay for screening compounds which affect transport and/or passive diffusion of proteins across the nuclear envelope in mammalian cells, which comprises the steps of:

a) treating a stably transfected cell line of the present invention with a compound; and

b) visualizing the distribution of reporter protein in nuclei and cytoplasm.

In accordance with another embodiment of the present invention, there is provided an *in vitro* assay for screening compounds which affect transport and/or passive diffusion of proteins across the nuclear envelope in mammalian cells, which comprises the steps of:

a) treating semi-permeabilized cells which support *in vitro* nuclear transport with a compound and a protein of the present invention; and

b) visualizing the distribution of reporter protein in nuclei and cytoplasm.

In accordance with one embodiment of the present invention, the cells may be treated with the compound and protein concurrently or one after the other.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates a schematic representation of different fluorescent reporter proteins containing GFP used in this study;

Fig. 2 illustrates transient expression of fusion proteins harboring one copy of GFP;

Fig. 3 illustrates transient expression of transport substrates with two and four copies of GFP;

Fig. 4 illustrates the effect of linker sequences on the localization of SV40-GFP4;

Fig. 5 illustrates the localization of reporter proteins harboring wild type and mutant versions of SV40-NLS;

Figs. 6A and 6B illustrate nucleocytoplasmic diffusion of NP-GFP, (A) chilling of cells, (B) treatment with sodium azide/deoxyglucose at different temperatures;

Fig. 7 illustrates the kinetics of the diffusion of NP-GFP across the nuclear envelope in HeLa cells treated with sodium azide/deoxyglucose; and

Fig. 8 illustrates stably transfected HeLa cells expressing nucleoplasmin-GFP.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the present invention, there is provided an experimental system to study nucleocytoplasmic traffic of proteins in living mammalian cells. Towards this goal, substrates were generated that contain one or more copies of *Aequorea victoria* green fluorescent protein (GFP). Reporter proteins of various sizes that carry different nuclear localization sequences (NLSs) were created to follow facilitated transport and passive diffusion across the nuclear envelope. The expression of reporter genes was controlled by an inducible promoter. Transiently and stably transfected HeLa cells were employed to follow the sorting of fluorescent reporter proteins. When NLS-GFP fusions were located in HeLa cells it was found that direct fusion of the NLS derived from SV40 T-anti-

gen to GFP prevented nuclear accumulation of the protein. However, insertion between NLS and GFP of different linkers encoding small amino acid residues produced reporter proteins that were competent for nuclear import.

Furthermore, it is demonstrated in accordance with the present invention that a small protein of approximately 30kD that harbors one copy of GFP and the bipartite *Xenopus* nucleoplasmin NLS accumulates efficiently in nuclei of HeLa cells. Under conditions which abolish active facilitated nuclear transport the protein diffuses across the nuclear envelope. This polypeptide can therefore be used as a reporter to analyze passive diffusion between nucleus and cytoplasm.

In addition to transient expression of novel transport substrates, stable cell lines that carry different genes for a fluorescent nuclear reporter protein were created. These cell lines provide an unique and novel tool for the characterization of nuclear protein import and diffusion in dividing human cells.

Based on the substrates to study nuclear traffic in growing cells of the present invention, histidine tagged versions of reporter proteins which permits their purification by metal affinity chromatography were also generated.

EXPERIMENTAL PROCEDURES

Cell culture and transfection

HtTA-1 cells are described in detail in Damke et al., 1995 (Damke, H. et al., 1995, *Meth. Enzym.*, **257**:209-220). HtTA-1 cells contain the tTA-element which permits the control of gene expression by addition of tetracycline to the growth medium.

HeLa cells were grown in Dulbecco's modified essential medium (DMEM) containing penicillin G (50U/ml), streptomycin (50µg/ml) and 8 % fetal bovine serum (complete medium). Cultures were maintained at 37°C in a humidified atmosphere containing 5% CO_2. For transient transfections cells were subcultured on six-well tissue culture plates and DNA was introduced with the calcium phosphate/DNA coprecipitation method and CalPhos Maximizer™ (Bio/Can, Mississauga, Ontario) or by use of the DOSPER™ Liposomal Transfection Reagent (Boehringer Mannheim, Canada) according to the manufacturer's instructions. Approximately 16 hrs to 24 hrs after addition of DNA, medium was removed and HeLa cells were kept in complete medium containing 50 nM dexamethasone. To generate stable cell lines, HtTA-1 cells were transfected following the modified transfection protocol described (Sambrook, J., Fritsch, E.F. and Maniatis, T., 1989, in *Molecular Cloning*, 2nd ed., pp 16.33-16.36, Cold Spring Harbor Laboratory Press).

Detection of transiently transfected cells expressing GFP-derivatives

Cells were incubated overnight with DNA followed by addition of fresh medium. Approximately 7 hours after the change of medium, cells were harvested by trypsinization and transferred to LabTek tissue culture chamber slides (Gibco, BRL). HeLa cells were kept for another 16 to 44 hours in complete medium containing 50 nM dexamethasone. To study diffusion, cells were washed with medium lacking dexamethasone and incubated in dexamethasone-free medium for 1 hour at 37°C prior to analysis of diffusion. To follow diffusion of nuclear substrates across the nuclear envelope, cells were washed with PBS and incubated in PBS in the cold. Alternatively, cells were treated with 10 mM sodium

azide and 10 mM 2-deoxy D-glucose in PBS as described for Figs. 6 and 7.

All subsequent steps were carried out at room temperature. For fixation slides were washed twice in phosphate buffered saline (PBS), incubated for 25 min in PBS/3.7% formaldehyde and washed three times with PBS. Samples were incubated with PBS containing 4',6'-Diamidino-2-phenylindole (DAPI) and 2 mg/ml bovine serum albumin (BSA), and washed twice in PBS. Slides were mounted in 30% glycerol in PBS. Slides were sealed with rubber cement and kept at 4°C. The fluorescence signals were stable for several weeks.

Generation of stable cell lines and detection of nuclear reporter proteins

HtTA-1 cells were grown and transfected as described above. Transfected cells were grown under standard conditions for at least two weeks without selection. One day after addition of 50 nM dexamethasone, cells were sorted by FACS into 24-well plates under sterile conditions. A maximum of one cell per well was obtained after sorting. Cells were further grown for two to three weeks and analyzed for the synthesis of nuclear reporter proteins. Towards this end, stable cell lines were kept in the presence of 50 nM dexamethasone (2 days) and analyzed by fluorescence microscopy as described above.

Plasmid constructions

Plasmid pGRE5-2, was used as a vector for genes encoding different fluorescent substrates. Gene expression is controlled by glucocorticoid response elements and can be induced by dexamethasone (Mader, S and White, J., 1993, *Proc. Natl. Acad. Sci USA*, **90**:5603-5607). A mutant version of GFP in pBluescriptKS$^+$ carrying a Ser65 --> Thr65 substitution by PCR-amplification of the GFP-DNA using appropriate primers was generated.

This created plasmid p580 which contains the mutant GFP allele cloned into the KpnI/EcoRI sites of pBluescript KS^+ (Chatterjee, S. and Stochaj, U., 1996, *BioTechniques*, **21**:62-63).

Plasmids encoding nuclear substrates with one or two copies of GFP

The generation of the gene encoding Nucleoplasmin-GFP, i.e., the C-terminal half of *Xenopus* nucleoplasmin fused to a single copy of GFP has been described (Chatterjee, S. and Stochaj, U., 1996, *BioTechniques*, **21**: 62-63). A gene fusion encoding nucleoplasmin NLS and GFP, termed NP-GFP, was created by fusion of oligonucleotides encoding the NLS (Barth, W. and Stochaj, U., 1996, *Biochem. Cell Biol.*, **74**: 363-372) to GFP. Plasmid p709 is derived from pGRE5-2 and codes for NP-GFP. Gene fusions containing two copies of GFP were created by removal of the UAG stop codon of GFP followed by insertion of a NotI linker (12mer, New England Biolabs). In addition, a NotI-linker was inserted into the KpnI-site of plasmid p580. Both copies of the GFP coding sequence were fused in frame via their NotI-sites thereby creating GFP2 in pBluescript (plasmid p690). Insertion of the KpnI/EcoRI fragment of p690 into the KpnI/EcoRI sites of pGRE5-2 created plasmid p785. For p785, the expression of GFP2 is driven by an inducible promoter carrying 5 glucocorticoid response elements.

Plasmids encoding fusion proteins containing four copies of GFP

DNAs encoding different NLSs were cloned into pGRE5-2. These vectors were digested with XhoI and EcoRI and a 2.8 kb XhoI/EcoRI fragment encoding four copies of GFP was ligated into the vector DNA. Following this ligation, DNA was digested with XhoI and AgeI. Recessed ends were filled in and DNA was autoli-

gated. The resulting plasmids p732, p733, p734 and p767 encode fusion proteins NP-GFP4, SV40-GFP4, SV40*-GFP4, and SV40inv-GFP4, respectively (Fig. 1). Derivatives of plasmids carrying wild type or mutant versions of SV40-NLS fused to GFP4 were generated by in frame insertion of linker sequences. Constructs encoding a direct fusion of NLS and GFP were digested with AgeI, incubated with T4-polymerase in the presence of 4 dNTPs. 8mer linkers (New England Biolabs) carrying a BamHI, EcoRI or SmaI-site were inserted. Linker insertion generated the following additional amino acid residues between NLS and the first copy of GFP: BamHI ala-asp-pro-pro (ADPP), EcoRI gly-asn-ser-pro (GNSP), SmaI ala-pro-gly-pro (APGP). Protein SV40-GFP4 that contains the linker sequence ala-asp-pro-pro is referred to as SV40-ADPP-GFP4. All other constructs are named accordingly. Plasmid p713 encodes GFP4 whose expression is controlled by five glucocorticoid response elements. To obtain this plasmid a KpnI/EcoRI fragment coding for the in frame fusion of four copies of GFP was inserted into the KpnI/EcoRI sites of pGRE5-2.

Plasmids to generate gene fusions containing GFP or NP-GFP

To facilitate the generation of gene fusions containing different forms of GFP the following constructs were generated: Plasmid p827 was obtained by inserting a XbaI/PstI fragment encoding a mutant GFP (L64 T65) into the XbaI/PstI sites of pKS^+.

p657 is derived from p580; p580 was digested with BsmI and treated with T4-DNA-polymerase in the presence of 4 dNTPs. A NotI-linker (12mer, New England Biolabs) was ligated to the filled in DNA, followed by digestion with NotI and autoligation. p657 is used to create fusion genes that encode GFP at their 5'-end. A derivative of p657 was generated by inserting the

NcoI/HpaI fragment of p827 into p657 that was treated with NcoI and HpaI. The resulting plasmid p833 contains GFP (L64 T65) carrying a NotI-linker at the former BsmI-site.

Similarly, p834 was created by insertion of the NcoI/HpaI fragment of p827 into plasmid p580.

Plasmids encoding histidine-tagged versions of fluorescent marker proteins

To generate genes encoding histidine-tagged marker proteins, coding sequences for the various fluorescent substrates were inserted in the appropriate pQE vector (Qiagen) to create in frame fusions of the histidine tag and the reporter protein. Specifically, to obtain His6-GFP4, plasmid p713 was digested with AgeI and EcoRI, blunted and inserted into pQE11 which was treated with BamHI and Klenow polymerase. These manipulations created plasmid p779 which encodes His6-GFP4 under control of the *Escherichia coli* tac promoter. Plasmid p779 was digested with NotI, the 4.8kb fragment was autoligated resulting in plasmid p788 which encodes His6-GFP2. Treatment of plasmid p779 with HpaI and autoligation of the 4.1kb linearized band generated plasmid p790 which carries the coding sequence for His6-GFP1. To generate His6-NP-GFP4, a HindIII/EcoRI fragment containing the coding sequence for NP-GFP4 was blunted and cloned into pQE10 treated with BamHI and Klenow-polymerase to remove recessive ends. Plasmid p778 codes for protein His6-NP-GFP4. Treatment of p778 with NotI and autoligation of the 4.8kb band generated plasmid p791 which encodes His6-NP-GFP2. Digestion of plasmid p778 with HpaI and autoligation of the 4.1kb band results in plasmid p792 which codes for His6-NP-GFP. To create histidine-tagged versions of GFP4 derivatives carrying different SV40-NLSs, the respective pGRE5-2 vectors containing the

constructs described were digested with EcoRI and HindIII followed by treatment with T4 DNA-polymerase in the presence of 4 dNTPs. Purified 2.8kb fragments were inserted in pQE10 which had been treated with BamHI and T4 DNA-polymerase plus 4 dNTPs. Plasmids p778, p791, p792, p799, p788, p790, p818, p819, p824, p828, p829 and p830 and the different NLSs they encode are listed in Table 1.

The different sequences listed in Table 1 are described in Barth and Stochaj, 1996 (Barth, W. and Stochaj, U., 1996, *Biochem. Cell Biol.*, **74**:363-372). Inserted linker sequences BamHI, EcoRI and SmaI are shown in italics (Table 1). Plasmids p827, p833 and p834 carry a mutant GFP (F64L S65T) as described by Cormack et al., 1996 (Cormack, B.P. et al., 1996, *Gene*, **174**: 33-38). The backbones of the plasmids and the respective fusion genes are listed in Table 1. In addition, it is stated which organisms can be used for the inducible expression of gene fusions.

Table 1

Generation of different plasmids to analyze nuclear transport of proteins and diffusion across the nuclear envelope

Plasmid name	Backbone	Gene encoded	Inducible expression in
p580	pBluescript KS+	GFP	---
P585	pGRE5-2	Nucleoplasmin-GFP	mammalian cells
p657	pBluescript KS+	GFP carrying a NotI-linker at 3'-end	---
p827	pBluescript KS+	GFP L64 T65	---
p833	pBluescript KS+	same as p657 but GFP L64 T65	---
p834	pBluescript KS+	same as p657 but GFP L64 T65	---
P709	pGRE5-2	NP-GFP	mammalian cells

p775	pBluescript KS+	NP-GFP	---
p776	pBluescript KS+	NP-GFP (blue GFP)	---
p777	pGRE5-2	NP-GFP (blue GFP)	mammalian cells
p732	pGRE5-2	NP-GFP4	mammalian cells
p733	pGRE5-2	SV40-GFP4	mammalian cells
p734	pGRE5-2	SV40*-GFP4	mammalian cells
p767	pGRE5-2	SV40inv-GFP4	mammalian cells
p764	pGRE5-2	SV40-*SmaI*-GFP4	mammalian cells
p765	pGRE5-22	SV40-*EcoRI*-GFP4	mammalian cells
p795	pGRE5-2	SV40-*BamHI*-GFP4	mammalian cells
p780	pGRE5-2	SV40inv-*BamHI*-GFP4	mammalian cells
p781	pGRE5-2	SV40*-*BamHI*-GFP4	mammalian cells
p782	pGRE5-2	SV40inv-*SmaI*-GFP4	mammalian cells
p783	pGRE5-2	SV40*-*SmaI*-GFP4	mammalian cells
p784	pGRE5-2	SV40inv-*EcoRI*-GFP4	mammalian cells
p785	pGRE5-2	GFP2	mammalian cells
p778	pQE	His6-NP-GFP4	*Escherichia coli*
p791	pQE	His6-NP-GFP2	*Escherichia coli*
p792	pQE	His6-NP-GFP	*Escherichia coli*
p779	pQE	His6-GFP4	*Escherichia coli*
p788	pQE	His6-GFP2	*Escherichia coli*
p790	pQE	His6-GFP	*Escherichia coli*
p818	pQE	His6-SV40-*BamHI*-GFP4	*Escherichia coli*
p819	pQE	His6-SV40inv-*BamHI*-GFP4	*Escherichia coli*
p824	pQE	His6-SV40*-*BamHI*-GFP4	*Escherichia coli*
p828	pQE	His6-SV40-*SmaI*-GFP4	*Escherichia coli*
p829	pQE	His6-SV40*-*SmaI*-GFP4	*Escherichia coli*
p830	pQE	His6-SV40inv-*SmaI*-GFP4	*Escherichia coli*

Plasmids encoding different mutants of GFP

In accordance with the present invention, additional plasmids which are identical to those listed in

Table 1 but carry a different version of GFP are currently being generated. GFP used to generate these plasmids carries a mutation at position at amino acid residue 64. The wild type Phe64 was replaced by Leu64 (F64L). This mutant GFP-derivative GFPmut1, referred to as GFP L64 T65, was obtained from B. R. Cormack and published in 1996 (*Gene*, **174**:33-38). In addition, nuclear substrates carrying the blue version of GFP are being created. The blue version of GFP is mutant P4 obtained from Roger Tsien (Heim, R. et al., 1994, *Proc. Natl. Acad. Sci.*, USA, **91**:12501-12504).

Plasmids for expression in *Saccharomyces cerevisiae*

Genes encoding different nuclear substrates described in Table 1 are presently cloned into yeast expression vectors. The expression of gene fusions in these vectors is driven by the *GAL1* control region. Gene expression is therefore induced when cells are grown on galactose as carbon source.

USE OF THE INVENTION

Generation of nuclear substrates containing GFP

Different genes encoding fusion proteins that are substrates for nuclear transport were created. The various constructs used in our studies are depicted in Fig. 1. Controls include proteins with one, two and four copies of GFP, which are termed GFP, GFP2 and GFP4, respectively. Fusion proteins were generated that carry the nucleoplasmin bipartite NLS (NP), the C-terminal half of nucleoplasmin (nucleoplasmin), SV40 wild type NLS (SV40), a mutant derivative (SV40*) and the inverse signal (SV40inv). Additional fusions were created that contain linker sequences between the NLS and the first copy of GFP. Only one construct harboring a linker sequence L and SV40-NLS is shown as an example. The different linker sequences inserted and

fusion proteins containing these linkers are described below.

Control proteins include GFP and a fusion protein containing two and four copies of GFP, termed GFP2 and GFP4, respectively. Polypeptide Nucleoplasmin-GFP contains the C-terminal half of *Xenopus* nucleoplasmin including the bipartite NLS and the DNA-binding region. Proteins SV40-GFP4, SV40*-GFP4, SV40inv-GFP4 and NP-GFP4 carry wild type or mutant forms of the NLS at their N-termini as indicated. Derivatives of the proteins containing different linker sequences between NLS and GFP4 are described above. The estimated molecular masses of GFP2 is about 60kD, for GFP4 and the NLS-containing derivatives approximately 120kD. When synthesized in *E. coli* and analyzed by denaturing gels, proteins migrate as expected.

Transient expression of GFP-containing transport substrates in HeLa cells

To determine the cellular localization of transport substrates, HeLa cells were transiently transfected with plasmids that express fusion proteins under the control of an inducible promoter. Gene expression was induced by addition of dexamethasone to the growth medium. As shown in Fig. 2 for plasmids encoding NP-GFP and GFP, in the presence of dexamethasone cells exhibit bright fluorescence. HeLa cells were transiently transfected with plasmids encoding NP-GFP (panels A, B) or GFP (panels C, D) and gene expression was induced with dexamethasone as described herein. DAPI-staining of the DNA (panels A, C) and green fluorescence (panels B, D) is shown. NP-GFP is efficiently targeted to the nucleus due to the presence of the bipartite NLS derived from *Xenopus* nucleoplasmin. By contrast, non-transfected cells display only a low level of autofluorescence (Fig. 2). Low autofluorescence signals were also

obtained for non-induced cells, i.e., cells kept in the absence of dexamethasone.

When GFP lacking a NLS was synthesized in HeLa cells, GFP was detected in both the cytoplasm and the nucleus (Fig. 2). This distribution is expected, since the small size of the protein permits its translocation across the nuclear envelope by passive diffusion. We also expressed a gene encoding two copies of GFP (GFP2) in HeLa cells. GFP2, a protein with a molecular mass of approximately 60kD was also found to traverse the nuclear envelope (Fig. 3). HeLa cells synthesizing GFP2 (panels A, B), GFP4 (panels C, D) and NP-GFP4 (panels E, F) were transiently transfected and processed in accordance with the present invention. DAPI-staining (panels A, C, E) and green fluorescence (panels B, D, F) are illustrated. The amount of GFP2 in nuclei, however, was relatively low, most likely the larger molecular mass restricted its diffusion across the nuclear pore.

The nuclear pore mediates transport of macromolecules across the nuclear envelope and molecules with a molecular mass of more than 70kD are excluded from the nucleus unless they carry a NLS. To obtain reporter proteins that exceed the size for passive diffusion, we have generated substrates that contain four copies of GFP. As shown in Fig. 3, in transiently transfected cells GFP4 was restricted to the cytoplasm and could not be detected in nuclei (panels C, D). On the basis of these results we created several different reporter proteins that contain four copies of GFP in addition to wild type or mutant forms of NLSs (Fig. 1).

Our initial experiments included the generation of constructs for which the NLS was directly fused to the first copy of GFP. For the bipartite nucleoplasmin-NLS, i.e., NP-GFP4, efficient import into nuclei was

obtained (Fig. 3, panels E, F). By contrast, constructs containing wild type SV40-NLS fused to GFP4 did not accumulate in nuclei (Fig. 4, panels A and B). Similarly, reporter proteins containing mutant versions of SV40-NLS fused to GFP4 were excluded from nuclei. It has previously been reported for SV40-NLS and other NLSs that flanking protein sequences, i.e., context of the NLS can affect the efficiency of nuclear targeting. We therefore reasoned that addition of amino acid residues between SV40-NLS and GFP4 could improve nuclear targeting of the fusion protein. Different linker sequences encoding short stretches of amino acid residues were inserted between SV40-NLS and GFP4. Linker sequences coded for proline and amino acid residues with small side chains that have the potential to form β-turns. Residues with small side chains are expected to provide flexibility of the linker. The linkers used generated the following extra amino acid sequences (one letter code) BamHI - ADPP, EcoRI - GNSP and SmaI-APGP. Reporter proteins containing wild type SV40-NLS and one of the different linker sequences were found to accumulate in nuclei (Fig. 4, panels C-H). Fusion proteins SV40-GFP4 containing SV40-NLS directly fused to GFP (panel A, B), or connected by different linker sequences were localized. Insertion of the peptide sequence ADPP (panels C, D), GNSP (panels E, F) and APGP (panels G, H) resulted in nuclear localization of reporter proteins. DAPI staining (panels A, C, E, G) and green fluorescence (panels B, D, F, H) were analyzed in parallel.

This demonstrates that linker insertion generated altered nuclear substrates that are recognized by the nuclear transport apparatus and targeted to the nucleus. We concentrated on constructs containing linkers ADPP or APGP for further studies. Towards this end,

linker insertions were created for SV40*-GFP4 and SV40inv-GFP4, and we analyzed the cellular localization of the respective proteins (Fig. 5). GFP4-derivatives containing the linker sequence ADPP (Panels A to F) or APGP (panels G to L) were transiently synthesized in HeLa cells. Substrate proteins contained wild type SV40-NLS (panels A, B, G, H), mutant SV40*-NLS (panels C, D, I, J) or inverse SV40inv (panels E, F, K, L). Staining with DAPI (panels A, C, E, G, I, K) and green fluorescence (panels B, D, F, H, J, L) is shown.

Transport substrates carrying SV40inv-NLS were excluded from the nucleus (Fig. 5). By contrast, proteins containing the mutant SV40*-NLS were not excluded from nuclei. However, these reporter proteins did not accumulate in nuclei. These results are consistent with previously published data that show nuclear targeting via this mutant NLS albeit with reduced efficiency.

Taken together, we have generated several nuclear transport substrates containing four copies of GFP and different NLSs. The type and context of the NLS specifies the localization of the reporter protein that can be easily followed in living cells.

Diffusion from the nucleus into the cytoplasm

In order to study diffusion of a polypeptide across the nuclear envelope, a reporter protein is required that is small enough to traverse the nuclear pore in the absence of nuclear import or export signals. NP-GFP, i.e., a fusion protein that contains the *Xenopus* nucleoplasmin NLS fused to one copy of GFP accumulated in nuclei under normal growth conditions (Fig. 2). However, the small size of approximately 30kD should permit NP-GFP to also traverse the nuclear pore complex by passive diffusion. As described above, transiently transfected cells synthesizing NP-GFP showed bright nuclear fluorescence (Fig. 2). To test whether

this protein can diffuse across the nuclear envelope, the inducer dexamethasone was removed and cells were transferred to 4°C. Although this led to diffusion of NP-GFP from the nucleus into the cytoplasm, even after 4 hours of incubation at 4°C the nuclear fluorescence was more intense than the fluorescence signal seen for the cytoplasm. If cells were incubated for 4 hours at 4 °C and returned to 37°C in growth medium cytoplasmic NP-GFP was re-imported into nuclei demonstrating that active facilitated transport was regained under normal growth conditions. Prolonged incubation at 4°C resulted in rounding of HeLa cells which was also reversible. Round cells flattened again and spread on the slide after overnight incubation at 37°C.

A more pronounced exit of NP-GFP from the nucleus into the cytoplasm was obtained when cells were incubated 2 hours at 4°C and subsequently kept for 2 hours on ice. Under these conditions fluorescence signals observed for nuclei and for the cytoplasm were of similar intensity, indicating the efficient equilibration of NP-GFP between nucleus and cytoplasm (Fig. 6A). HeLa cells transiently expressing the NP-GFP gene were incubated for 2 hours at 4°C, followed by 2 hours on ice. The distribution of fluorescence was determined before (panels A, B) and after chilling (panels C, D). After incubation on ice cells were returned to 37°C and incubated overnight in the absence of the inducer dexamethasone to test for re-import of NP-GFP into the nucleus (panels E, F). Panels A, C, D show phase contrast, panels B, D and F the distribution of fluorescence.

These results support the idea that in chilled HeLa cells NP-GFP left the nucleus by diffusion and nuclear accumulation of NP-GFP is prevented due to the absence of active facilitated transport into the

nucleus. To determine whether nuclear exit of NP-GFP is reversible, cells were first treated 2 hours at 4°C and 2 hours on ice and subsequently incubated in growth medium at 37°C (Fig. 6A). Under these conditions cytoplasmic NP-GFP was re-imported into nuclei demonstrating that diffusion across the nuclear envelope was reversible and cells recovered from chilling.

In addition to cold treatment, we also tested the effect of sodium azide/2-deoxy D-glucose (azide/deoxyglucose) on the localization of NP-GFP. After incubation of transiently transfected cells with azide/deoxyglucose, NP-GFP equilibrated efficiently between nucleus and cytoplasm (Fig. 6B). Transiently transfected HeLa cells containing NP-GFP were incubated with azide/deoxyglucose as described herein. The localization of NP-GFP was determined before (panels A, B and G, H) and after treatment with sodium azide/deoxyglucose for 45 min (panels C, D and I, J). Slides were washed in PBS and incubated overnight with medium lacking dexamethasone to analyze accumulation of NP-GFP in the nucleus (panels E, F and K, L). Identical experiments were carried out on ice (panels A - F) and at room temperature (panels G - L). Phase contrast is presented in panels A, C, E, G, I, K, the location of fluorescent proteins is shown in panels B, D, F, H, J and L.

We further determined whether incubation with azide/deoxyglucose led to nuclear exit of NP-GFP at other temperatures. As demonstrated in Fig. 6B treatment at room temperature or at 37°C results in diffusion of NP-GFP across the nuclear envelope. For all the different temperatures tested, we found that upon return of cells to normal growth medium and 37°C, reimport of NP-GFP was observed for at least a fraction of cells. After incubation with azide/deoxyglucose at

37°C fewer cells seemed to recover when compared to those treated at room temperature or incubated on ic .

To study diffusion of NP-GFP across the nuclear envelope in more detail, we followed the distribution of fluorescence in cells that were incubated with azide/deoxyglucose at room temperature (Fig. 7). HeLa cells were incubated at room temperature with sodium azide/deoxyglucose and the nucleocytoplasmic distribution of the fluorescence signal was determined for 0 min, 10 min, 30 min, 45 min, 60 min and 90 min of the incubation with sodium azide/deoxyglucose as indicated in Fig. 7.

A gradual increase of cytoplasmic fluorescence was seen upon treatment with azide/deoxyglucose. The equilibration of NP-GFP was complete at approximately 45 min, and prolonged incubations with inhibitors did not lead to further changes.

Generation and characterization of a stable cell line expressing Nucleoplasmin-GFP

When induced with dexamethasone, transiently transfected cells show differences in the fluorescence which reflects different copy numbers of the plasmids introduced into each cell. In addition, transiently transfected cells loose the plasmid after prolonged culturing. Therefore, for many studies it is advantageous to have a uniform population of cells with comparable fluorescence. To this end, we created a stable cell line which carries the gene for nucleoplasmin-GFP. Cells were transfected and grown for two weeks without selection. Dexamethasone was added to the growth medium and cells displaying fluorescence were sorted by FACS. Under these conditions, approximately one in 10,000 cells gave a fluorescence signal. Single cells were isolated and further cultured. Characterization of stable transfectants demonstrated that the fluorescence

signal observed varied between different clones. This most likely reflects variable levels of gene expression due to differences in the integration sites of the plasmid. However, fluorescence signals for single cells derived from the same clone were comparable (Fig. 8, panels C, D). HeLa cells were transfected with a plasmid encoding Nucleoplasmin-GFP and stable transfectants were selected as described herein. Green fluorescence obtained after induction of gene expression with dexamethasone (panel B) or in the absence of dexamethasone (panel D) was determined. Nuclei were detected by staining with DAPI (panel A, C).

In the absence of the inducer dexamethasone, no fluorescence signal was obtained demonstrating the tight control of gene expression in this experimental system (Fig. 8).

Additional stable cell lines have been generated that express NP-GFP and GFP only. We are presently creating a larger number of these stable cell lines to obtain clones with improved fluorescence.

Previous and future use of plasmids and cell lines

We have generated a nuclear transport system that permits the analysis of nuclear import and passive diffusion of proteins from the nucleus into the cytoplasm. To achieve this goal, we have created a variety of fusion proteins that contain one, two or four copies of GFP.

Generation of fusion proteins containing GFP and a nuclear localization signal

Different plasmid vectors encoding NLS-GFP derivatives will be useful to create fusion proteins that are targeted to the nucleus. If the protein of interest has deleterious effects, it is desirable to have an experimental system that permits the controlled induction of genes. To this end, promoters regulated by

steroid hormone binding sites as described here for pGRE-derivatives will be optimal.

Analysis of nuclear transport in mammalian cell lines

Analysis of facilitated transport in HeLa cells requires substrates that are unable to enter the nucleus by diffusion. Therefore, we have designed fusion proteins whose size prevents them from diffusing across the nuclear pore. These nuclear substrates contain four copies of GFP, and in the absence of a functional NLS GFP4 was excluded from nuclei. By contrast, protein GFP2 which contains two copies of GFP was present in nuclei, indicating that GFP2 could diffuse across the nuclear envelope. The translocation of GFP2 into nuclei is not mediated by a cryptic NLS since such a signal would also be present in GFP4, a protein we have shown to be excluded from nuclei. Thus, although the molecular mass of GFP2 is close to the exclusion size of nuclear pores for passive diffusion, it can still traverse the nuclear envelope of HeLa cells.

Since GFP4 remained cytoplasmic in HeLa cells, we concentrated on the generation of reporter proteins that harbor four copies of GFP. Nucleoplasmin NLS provided a functional signal when directly fused to the GFP4-moiety, demonstrating that the context provided by GFP4 permits the recognition of the bipartite signal and nuclear accumulation of NP-GFP4. In contrast to nuclear targeting via nucleoplasmin NLS, direct fusion of SV40-NLS to GFP4 did not result in nuclear targeting of the fusion protein. These results can be interpreted as the failure of cytoplasmic factors to bind the simple SV40-NLS when fused to GFP4 due to steric hindrance. Alternatively, the flanking sequences provided by GFP could interfere with its function. It is currently believed that nuclear transport of proteins

harboring SV40-NLS and the bipartite nucleoplasmin-NLS are both recognized by importin-alpha during nuclear transport. Since nucleoplasmin NLS directly fused to GFP or GFP4 was efficiently targeted to the nucleus, it is reasonable to assume that the N-terminal sequence of GFP does not inhibit binding to cytoplasmic NLS-receptors in general. We therefore favor the interpretation that SV40-NLS is not readily accessible for recognition by nuclear transport factors when directly fused to GFP4. Introduction of linker sequences that encode a stretch of small amino acid residues, however, enabled SV40-NLS to target GFP4 to the nucleus. Control experiments were carried out with fusion proteins containing identical linker insertions and SV40*-NLS or SV40inv-NLS. SV40*-GFP4 did not accumulate and SV40inv-GFP4 was excluded from nuclei demonstrating that the inserted linker sequence does not function as a cryptic NLS. When compared to GFP4 or SV40inv-GFP carrying different linker insertions it was observed that SV40*-GFP was not excluded from nuclei. This result is in agreement with previous studies which showed that the mutant NLS of SV40-T antigen harboring a Lys128->Thr128 mutation can still function as a weak NLS.

Analysis of diffusion across the nuclear envelope

As previously described by us the small size of GFP permits this protein to enter the nucleus by passive diffusion. Fusion of the bipartite nucleoplasmin NLS to GFP efficiently mediated nuclear accumulation of the reporter protein. The small size of NP-GFP, i.e., approximately 30kD, made it an ideal tool to study diffusion from the nucleus into the cytoplasm.

For the experiments reported here, we demonstrated that NP-GFP exited the nucleus at low temperatures even in the absence of metabolic inhibitors. Migration of NP-GFP into the cytoplasm of HeLa cells

during chilling is not caused by irreversible damage of the cells or the nuclear envelope. HeLa cells survived this treatment, and the nuclear substrate was re-imported into nuclei after shifting to normal growth conditions. The rapid nuclear exit of NP-GFP during chilling enables us to study diffusion in the absence of metabolic inhibitors which are likely to interfere with a variety of cellular processes.

In addition to the analyses described above, NP-GFP will be useful as a tool to test how various drugs might interfere with diffusion across the nuclear envelope. The possible effect by these agents can be easily followed by fluorescence microscopy.

Stable cell lines expression fluorescent reporter protein

We also studied diffusion of the fluorescent reporter protein in the stable cell line synthesizing nucleoplasmin-GFP. However, we did not detect an equilibration of the fluorescence signal between nucleus and cytoplasm if cells were chilled or treated with sodium azide/deoxyglucose. Two possible hypotheses which are not mutually exclusive could explain this result: (i) The larger size of nucleoplasmin-GFP as compared to NP-GFP prevents its rapid diffusion across the nuclear envelope. (ii) Nucleoplasmin-GFP still contains the DNA-binding region of *Xenopus* nucleoplasmin, and binding to DNA could retain the reporter protein in the nucleus even in the absence of active facilitated transport across the nuclear membrane. This cell line will therefore be useful for studies that concentrate on the analysis of nuclear transport only.

Additional cell lines expressing NP-GFP or only GFP have also been generated, and we are currently creating additional cell lines which have improved fluorescence.

In general, all of the cell lines described will allow the controlled expression not only of the fluorescent reporter protein but also of additional genes of interest. Since the cells contain the tTA-element (Damke, H. et al., 1995, *Meth. Enzym.*, **257**:209-220), they can be used to regulate the synthesis of a protein of interest by addition of tetracycline to the growth medium. Genes encoded on a tetracycline response plasmid are repressed in the presence of tetracycline, whereas in the absence of tetracycline gene expression is induced.

Use of nuclear substrates in other mammalian cell lines and in the yeast *S. cerevisiae*

Nuclear substrates described in this study were also tested in other mammalian cell lines such as COS-7 and CHO cells. Similar results as described for HeLa cells were obtained, indicating that the reporter system we have developed is useful as a model to study nuclear transport in different mammalian cells. Taken together, we have established tools to study diffusion and signal-mediated transport across the nuclear envelope in higher eukaryotic cells. This sets the stage to analyze nucleocytoplasmic traffic in living mammalian cells in more detail.

Expression of fluorescent reporter proteins described here is currently also analyzed in the yeast *S. cerevisiae*.

Analysis of nuclear transport *in vitro*

Histidine-tagged (His6) versions of fluorescent reporter proteins have been created. Synthesis of these proteins is inducible in *Escherichia coli*. Fusion proteins are currently purified by affinity-purification and analyzed in *in vitro* nuclear transport assays. The advantage of these substrates is that they are easy to purify in large quantities and at low costs.

Previously used substrates for *in vitro* studies involved the synthesis and covalent coupling of NLS-peptides, an expensive and laborious procedure.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth, and as follows in the scope of the appended claims.

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. Genes encoding reporter proteins to monitor transport and passive diffusion of proteins across the nuclear envelope in mammalian cells, which comprises a protein carrying a nuclear localization sequence (NLS) fused to at least one copy of a reporter protein, wherein the reporter protein allows for the visualization of the reporter fusion protein.

2. The genes encoding proteins of claim 1, wherein the NLS is connected to the reporter protein via a linker sequence encoding small amino acid residues.

3. The genes encoding proteins of claim 2, wherein the NLS is derived from *Xenopus laevis* nucleoplasmin or present in the C-terminal portion of nucleoplasmin.

4. The genes encoding proteins of claim 3, wherein said reporter protein is a fluorescent protein.

5. The genes encoding proteins of claim 4, wherein said fluorescent protein is *Aequorea victoria* green fluorescent protein (GFP).

6. An expression vector to transiently and stably express reporter genes of claim 1, which comprises a gene encoding a protein carrying a nuclear localization sequence (NLS) fused to at least one copy of a reporter protein, wherein expression of the gene is under the control of an inducible promoter.

7. The expression vector of claim 6, wherein the NLS is connected to the reporter protein via a linker sequence encoding small amino acid residues.

8. The expression vector of claim 7, wherein the reporter protein is a fluorescent protein.

9. The expression vector of claim 8, wherein the fluorescent protein is *Aequorea victoria* green fluorescent protein (GFP).

10. A cell line stably expressing a nuclear reporter protein to monitor transport and passive diffusion of proteins across the nuclear envelope, which comprises a cell line transfected by the expression vector of claim 6.

11. The cell line of claim 10, wherein NLS is fused to reporter protein via a linker encoding small amino acid residue.

12. The cell line of claim 10, wherein said cell line is selected from the group consisting of HeLa, COS and CHO.

13. An *in vivo* assay for screening compounds which affect transport and/or passive diffusion of proteins across the nuclear envelope in mammalian cells, which comprises the steps of:

a) treating a cell line of claim 10 with a compound; and

b) visualizing the distribution of reporter protein in nuclei and cytoplasm.

14. An *in vitro* assay for screening compounds which affect transport and/or passive diffusion of proteins across the nuclear envelope in mammalian cells, which comprises the steps of:

a) treating semi-permeabilized cells which support *in vitro* nuclear transport with a compound and a protein of claim 1; and

b) visualizing the distribution of reporter protein in nuclei and cytoplasm.

15. The assay of claim 14, wherein said cells can be treated with the compound and protein concurrently or one after the other.

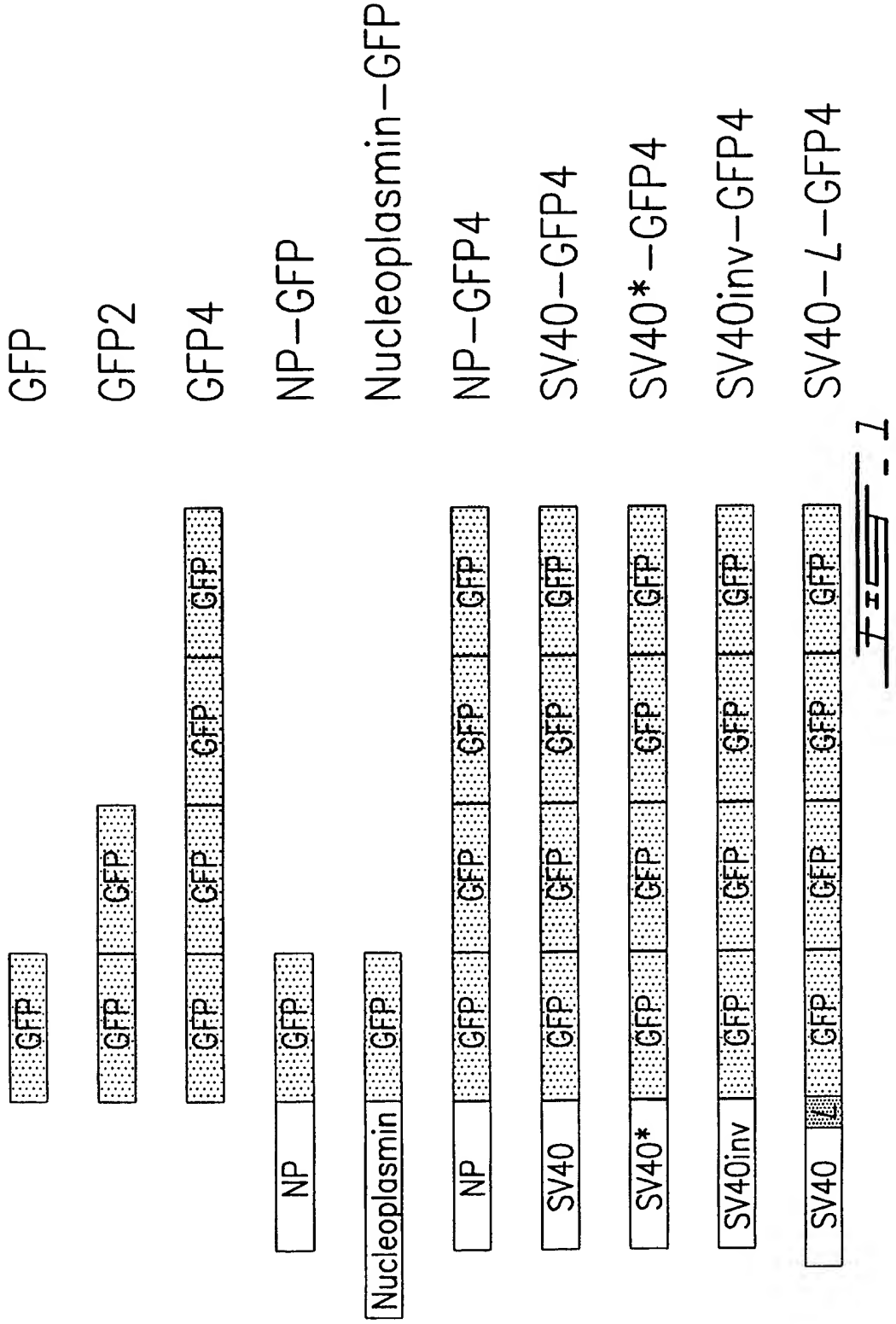
GFP
GFP2
GFP4
NP–GFP
Nucleoplasmin–GFP
NP–GFP4
SV40–GFP4
SV40*–GFP4
SV40inv–GFP4
SV40–L–GFP4
GFP
NP
Nucleoplasmin
SV40
SV40*
SV40inv
SV40

Fig. 1

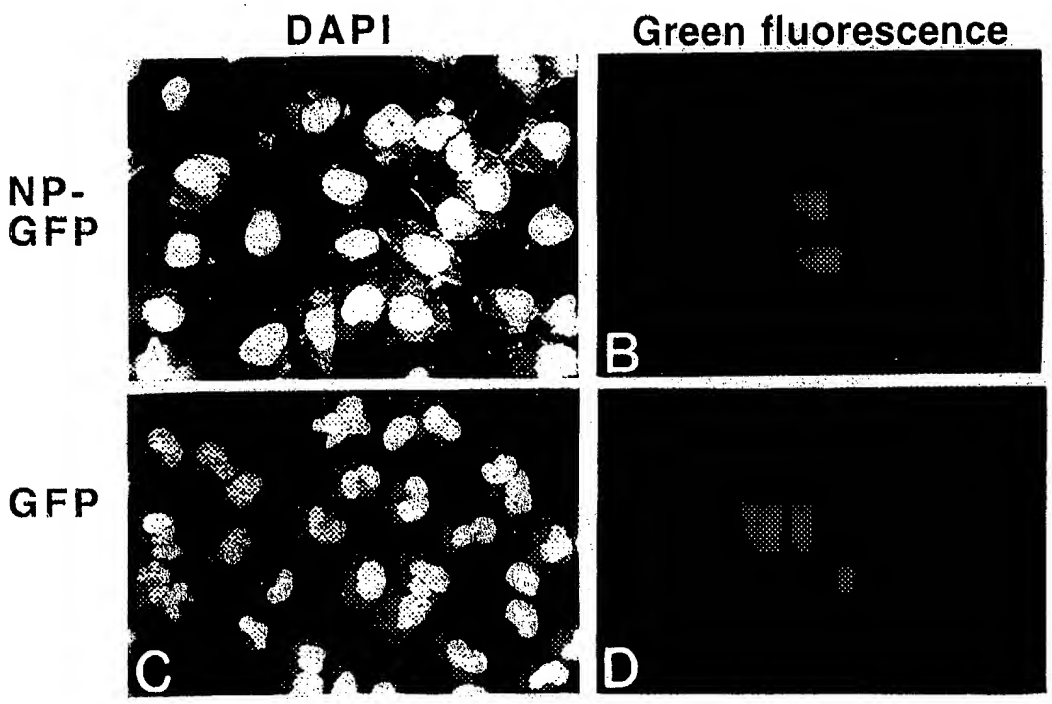

Fig. 2

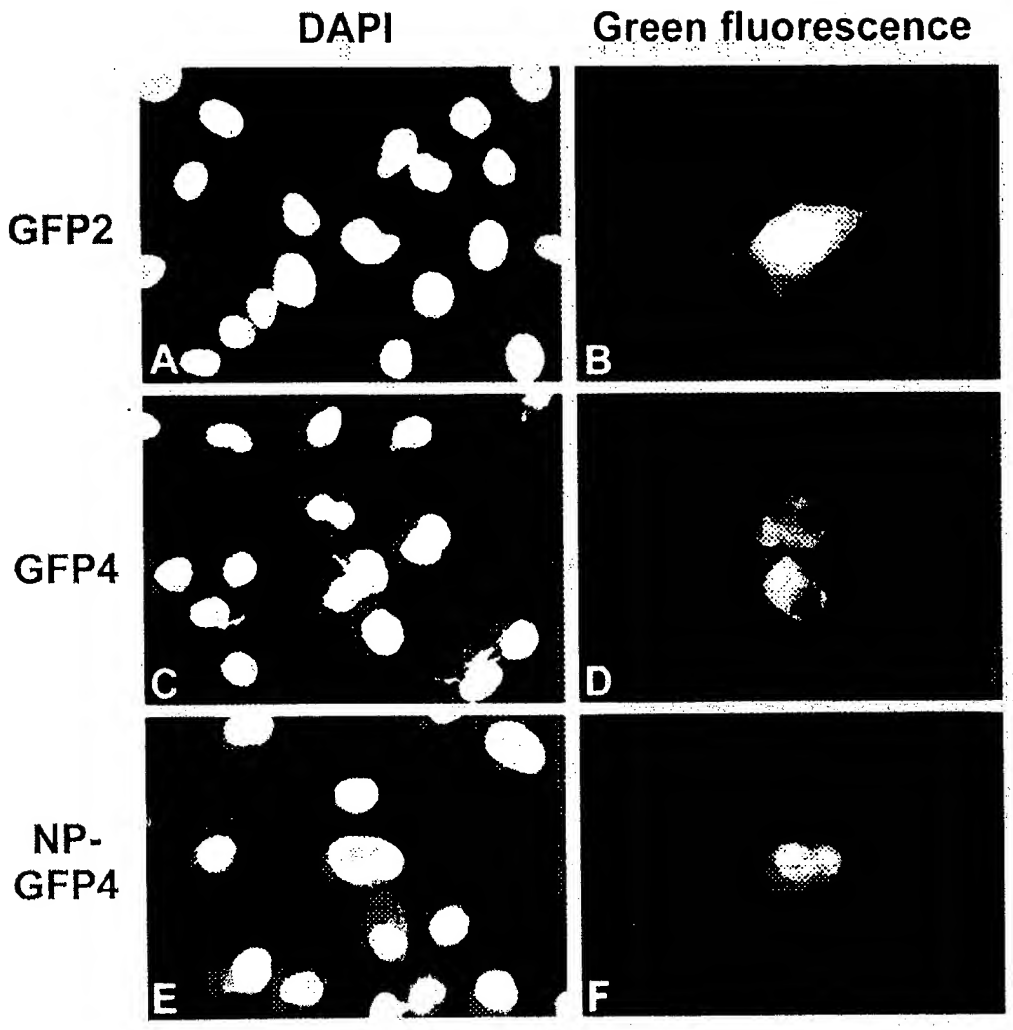

FIG. 3

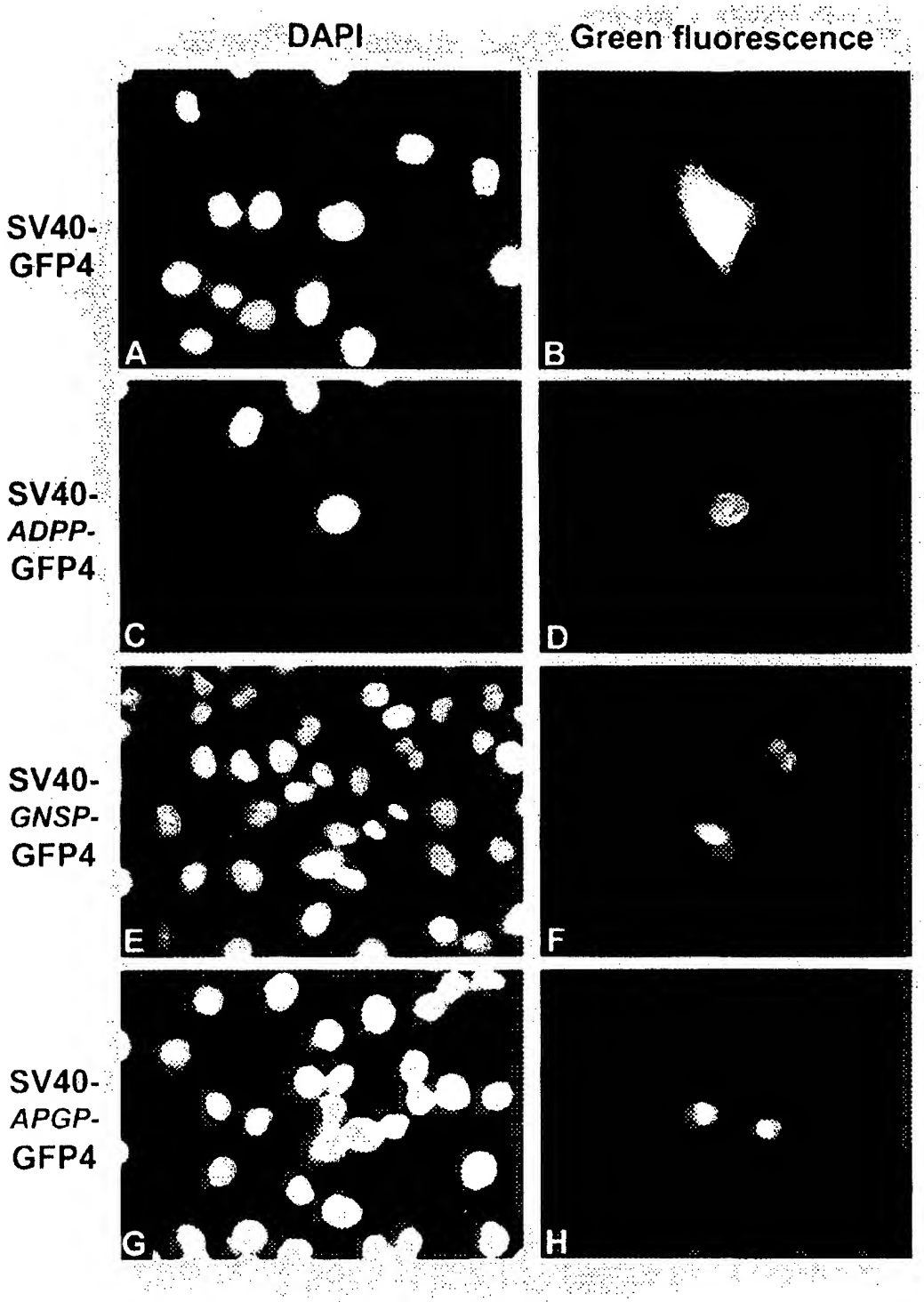
DAPI
Green fluorescence
SV40-GFP4
SV40-ADPP-GFP4
SV40-GNSP-GFP4
SV40-APGP-GFP4
A
B
C
D
E
F
G
H

FIG. 4

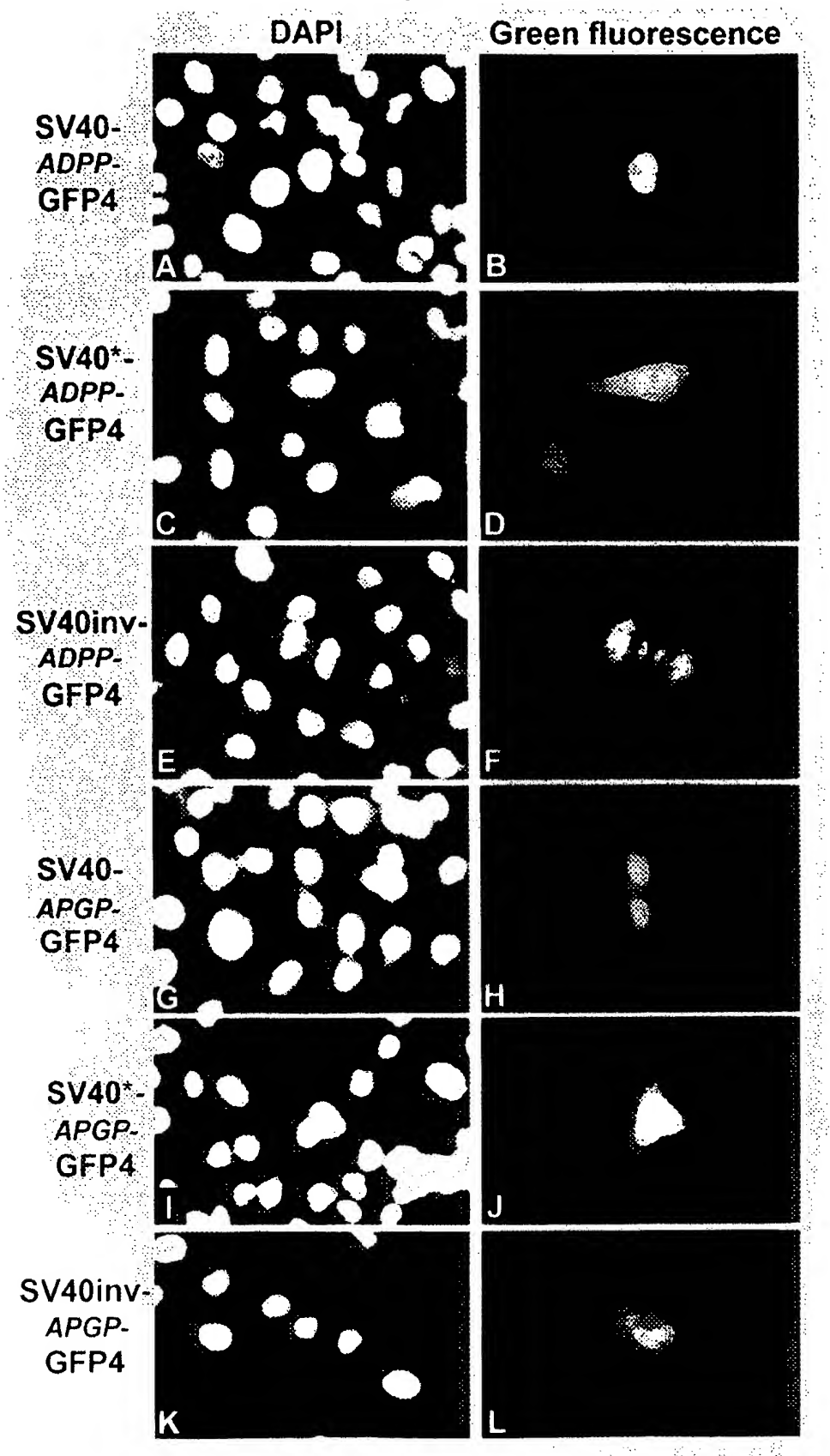
DAPI
Green fluorescence
SV40-
ADPP-
GFP4
A
B
SV40*-
ADPP-
GFP4
C
D
SV40inv-
ADPP-
GFP4
E
F
SV40-
APGP-
GFP4
G
H
SV40*-
APGP-
GFP4
I
J
SV40inv-
APGP-
GFP4
K
L

Fig. 5

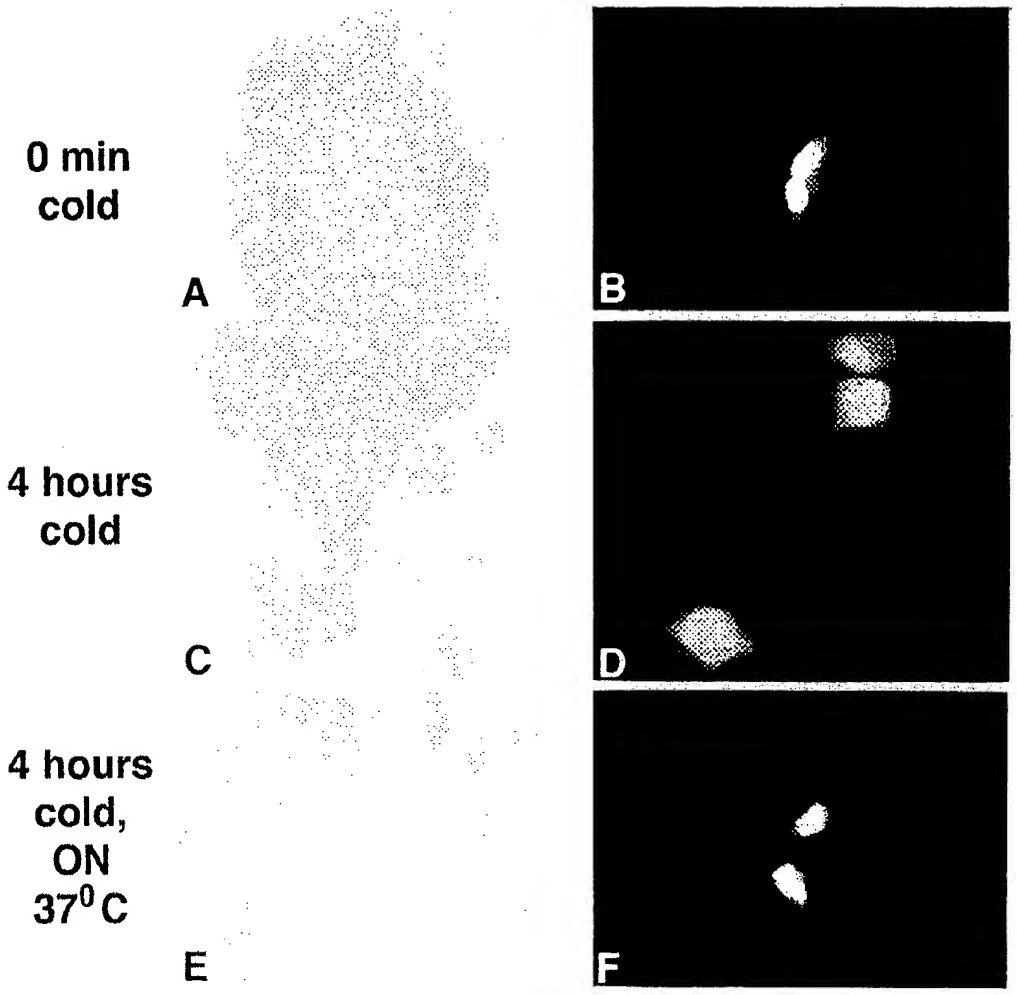

Fig. 6A

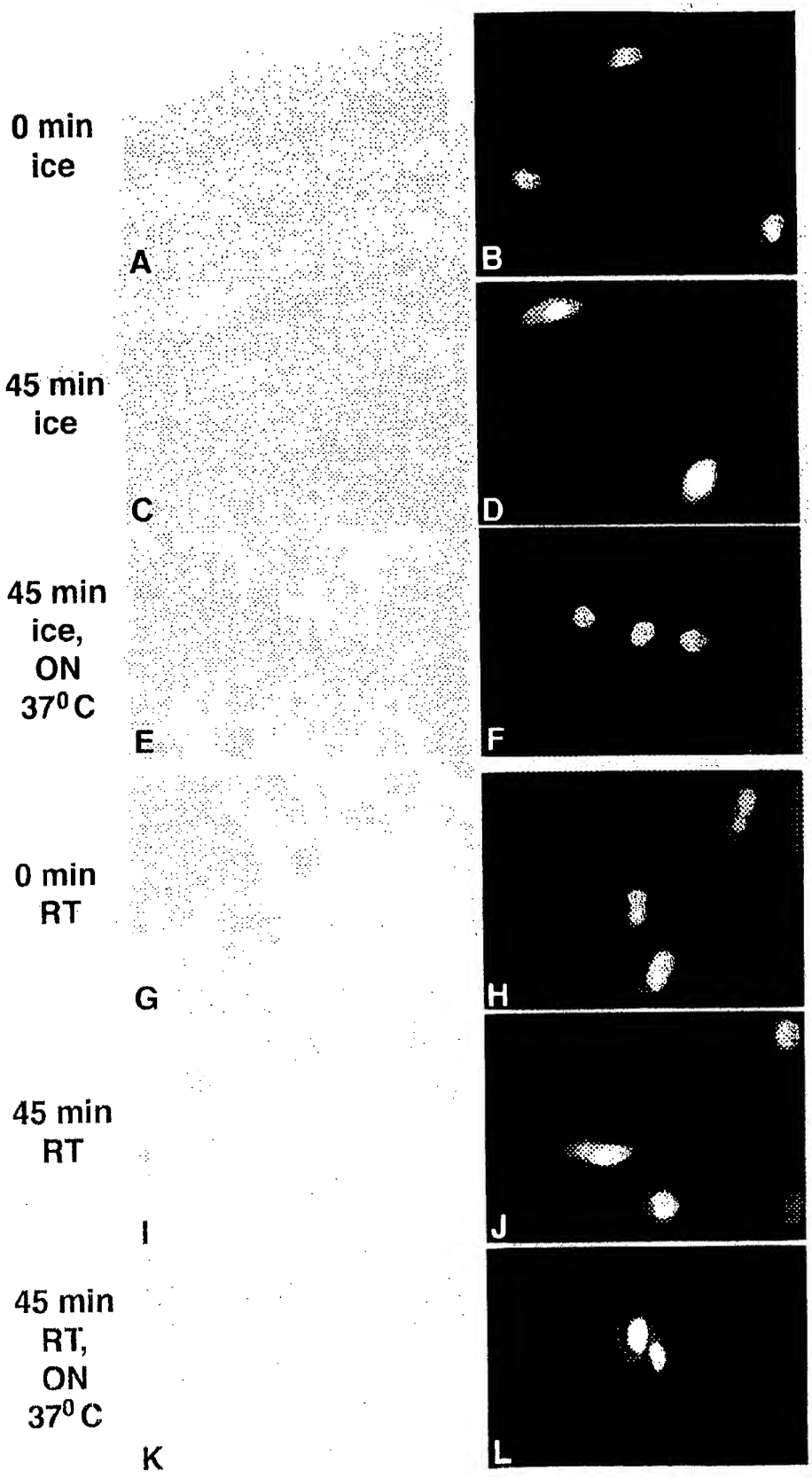
0 min ice
A
B
45 min ice
C
D
45 min ice, ON 37°C
E
F
0 min RT
G
H
45 min RT
I
J
45 min RT, ON 37°C
K
L

FIG 6B

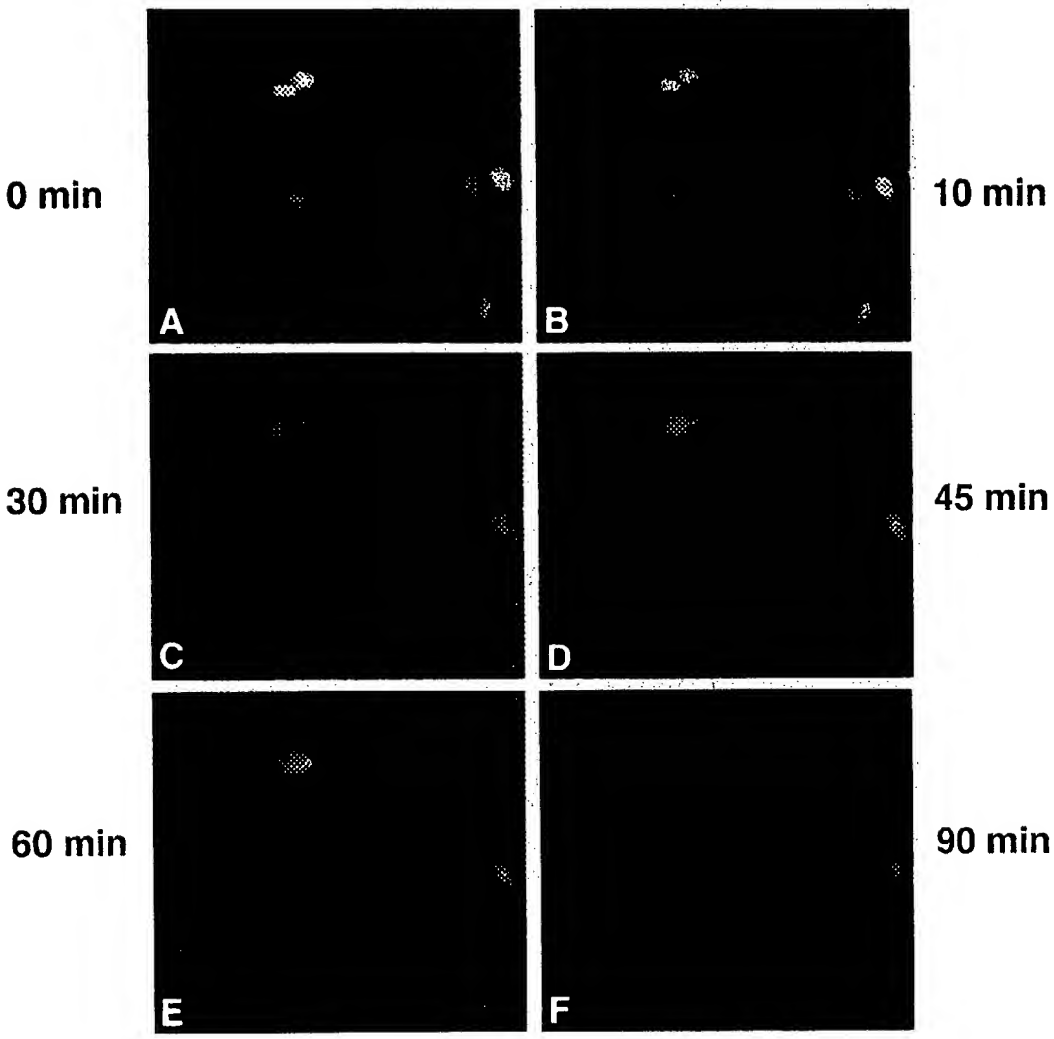

Fig. 7

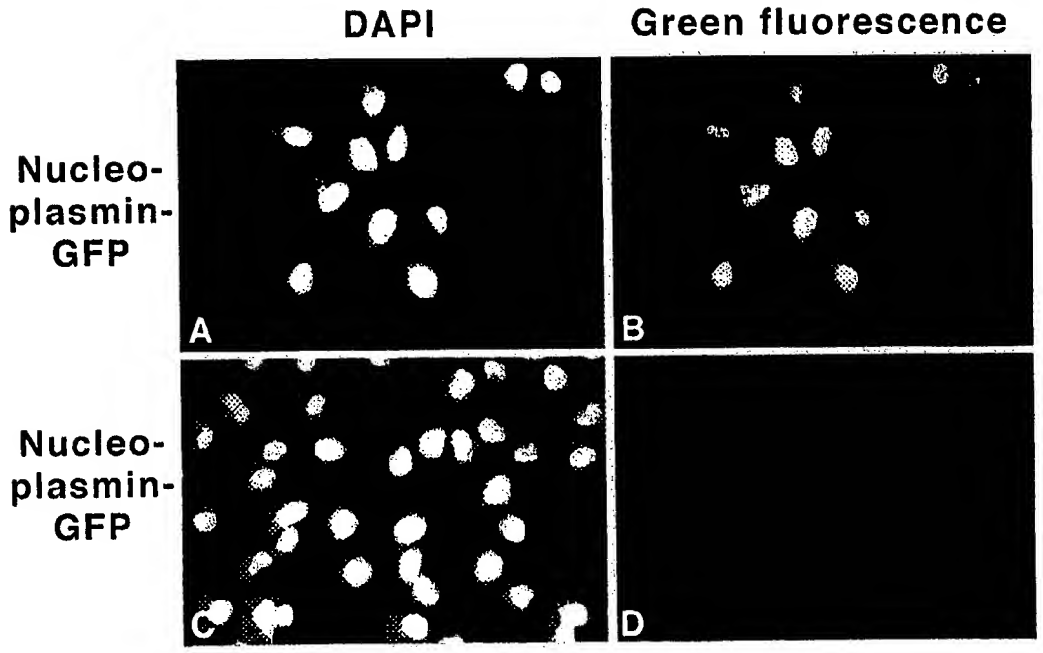
DAPI
Green fluorescence
Nucleo-plasmin-GFP
Nucleo-plasmin-GFP
A
B
C
D

Fig. 8

INTERNATIONAL SEARCH REPORT

Interr. ..nal Application No
PCT/CA 98/00385

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/62 C12N15/85 C12N5/10 G01N33/68

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SHIVANI CHATTERJEE ET AL.: "Monitoring nuclear transport in HeLa cells using the green fluorescent protein" BIOTECHNIQUES., vol. 21, no. 1, July 1996, pages 62-63, XP002076910 NATICK US cited in the application see the whole document --- -/--	1-15

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 September 1998	22/09/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Montero Lopez, B

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 98/00385

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	PTUSHKINA M ET AL: "Intracellular targeting and mRNA interactions of the eukaryotic translation initiation factor eIF4E in the yeast Saccharomyces cerevisiae." BIOCHIMICA ET BIOPHYSICA ACTA, (1996 AUG 14) 1308 (2) 142-50. JOURNAL CODE: AOW. ISSN: 0006-3002., XP002076911 Netherlands see abstract see page 145, right-hand column, paragraph 2 - page 148, right-hand column, paragraph 2 ---	1,2,4-15
X	WEIS K ET AL: "The conserved amino-terminal domain of hSRP1 alpha is essential for nuclear protein import." EMBO JOURNAL, (1996 APR 15) 15 (8) 1818-25. JOURNAL CODE: EMB. ISSN: 0261-4189., XP002076912 ENGLAND: United Kingdom see the whole document ---	1,2,6,7, 10-15
X	SCHREIBER V ET AL: "A eukaryotic expression vector for the study of nuclear localization signals." GENE, (1994 DEC 15) 150 (2) 411-2. JOURNAL CODE: FOP. ISSN: 0378-1119., XP002076913 Netherlands see the whole document ---	1,2,6,7, 10-15
X	WERNER BARTH ET AL.: "The yeast nucleoporin Nsp1 binds nuclear localization sequences in vitro" BIOCHEMISTRY AND CELL BIOLOGY, vol. 74, no. 3, 1996, pages 363-372, XP002076914 cited in the application see page 364, left-hand column, paragraph 5 see page 364, right-hand column, line 2 - line 3; figure 1 see page 368, right-hand column, paragraph 4 --- -/--	1,3,6, 10,12, 14,15

1

INTERNATIONAL SEARCH REPORT

Internal Application No
PCT/CA 98/00385

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	CHATTERJEE S ET AL: "In vivo analysis of nuclear protein traffic in mammalian cells." EXPERIMENTAL CELL RESEARCH, (1997 OCT 10) 236 (1) 346-50. JOURNAL CODE: EPB. ISSN: 0014-4827., XP002076915 United States see abstract see page 347, left-hand column, last paragraph - page 350, right-hand column, paragraph 1 ---	1-15
P,X	WO 97 41228 A (PIONEER HI-BRED INTERNATIONAL, INC.) 6 November 1997 see page 8, line 19 - line 34 see page 14, line 16 - line 27 see page 39, line 12 - page 22; example 6 -----	1,6,10

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte. onal Application No
PCT/CA 98/00385

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9741228 A	06-11-1997	AU 2998397 A	19-11-1997